AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2013JD020994

Key Points:
- To develop an objective approach to blend NLDAS drought indices
- To establish the linkage between USDM statistics and NLDAS drought indices
- To reconstruct long-term OBNDI

Correspondence to:
Y. Xia,
Youlong.Xia@noaa.gov

Citation:
Xia, Y., M. B. Ek, C. D. Peters-Lidard, D. Mocko, M. Svoboda, J. Sheffield, and E. F. Wood (2014), Application of USDM statistics in NLDAS-2: Optimal blended NLDAS drought index over the continental United States, *J. Geophys. Res. Atmos.*, *119*, 2947–2965, doi:10.1002/2013JD020994.

Received 7 OCT 2013
Accepted 28 JAN 2014
Accepted article online 1 FEB 2014
Published online 21 MAR 2014

Application of USDM statistics in NLDAS-2: Optimal blended NLDAS drought index over the continental United States

Youlong Xia[1,2], Michael B. Ek[1], Christa D. Peters-Lidard[3], David Mocko[3,4], Mark Svoboda[5], Justin Sheffield[6], and Eric F. Wood[6]

[1]Environmental Modeling Center, National Centers for Environmental Prediction, College Park, Maryland, USA, [2]IMSG at NCEP/EMC, College Park, Maryland, USA, [3]Hydrological Sciences Laboratory at Goddard Space Flight Center, National Aeronautics and Space Administration, Greenbelt, Maryland, USA, [4]SAIC, Greenbelt, Maryland, USA, [5]National Drought Mitigation Center, University of Nebraska–Lincoln, Lincoln, Nebraska, USA, [6]Department of Environmental and Civil Engineering, Princeton University, Princeton, New Jersey, USA

Abstract This study performs three experiments to calibrate the drought area percentages in the continental United States (CONUS), six U.S. Drought Monitor (USDM) regions, and 48 states downloaded from the USDM archive website. The corresponding three experiments are named CONUS, Region, and State, respectively. The data sets used in these experiments are from the North American Land Data Assimilation System Phase 2 (NLDAS-2). The main purpose is to develop an automated USDM-based approach to objectively generate and reconstruct USDM-style drought maps using NLDAS-2 data by mimicking 10 year (2000–2009) USDM statistics. The results show that State and Region have larger correlation coefficients and smaller root-mean-square error (RMSE) and bias than CONUS when compared to the drought area percentages derived from the USDM, indicating that State and Region perform better than CONUS. In general, State marginally outperforms Region in terms of RMSE, bias, and correlation. Analysis of normalized optimal weight coefficients shows that soil moisture percentiles (top 1 m and total column) play the dominant role in most of the 48 states. The optimal blended NLDAS drought index (OBNDI) has higher simulation skills (correlation coefficient and Nash-Sutcliffe efficiency) in the South, Southeast, High Plains, and Midwest regions when compared to those in the West and Northeast. The highest simulation skills appear in TX and OK. By using optimal equations, we can reconstruct the long-term drought area percentages and OBNDI over the continental United States for the entire period of the NLDAS-2 data sets (January 1979 to present).

1. Introduction

The multi-institution North American Land Data Assimilation project (NLDAS) has experienced four stages since it was initiated in 2000 [*Mitchell et al.*, 2004]. The first stage established infrastructure, including selection of land surface models, generation of surface forcing data, collection of soil and vegetation data sets, and in situ and satellite-retrieved observations. Four modeling groups ran their models for a 3 year period (from 1 October 1997 to 30 September 1999), separately. The National Centers for Environmental Prediction's (NCEP) Environmental Modeling Center ran the community Noah model, Princeton University's land group ran the VIC (Variable Infiltration Capacity) model, NASA Goddard Space Flight Center's hydrology group ran the Mosaic model, and the National Weather Service's Office of Hydrologic Development ran the SAC (Sacramento Soil Moisture Accounting) hydrological model. The model outputs were evaluated and compared with in situ observations and satellite-retrieved products. The overall results showed that all four models are able to capture broad features for these validated variables such as energy fluxes (e.g., net radiation, sensible heat, latent heat, and ground heat), water fluxes (i.e., evapotranspiration and total runoff), and state variables (i.e., soil temperature, soil moisture, land surface temperature, snow cover fraction, and snow water equivalent). The validation tools and overall results are detailed in *Mitchell et al.* [2004].

The second stage focused on improving model physics, tuning model parameters, and improving surface forcing data quality and reliability based on the findings from the first stage, and further expanding the short-term (i.e., 3 years) model products to long-term ($>$ 30 years) model products. The NCEP NLDAS team improved Noah simulations during the cold season [*Livneh et al.*, 2010] and warm season [*Wei et al.*, 2013]

through collaboration with the University of Washington. The Princeton Land group improved the VIC simulation by calibrating model parameters [*Troy et al*., 2008], and the NCEP NLDAS team also improved SAC simulations by using climatologically averaged observed potential evaporation [*Xia et al*., 2012a]. For surface forcing data, the CPC (Climate Prediction Center) gauge precipitation was bias corrected with the Parameter-elevation Regressions on Independent Slopes Model precipitation [*Daly et al*., 1994] through reducing the impact of topography on gauge precipitation. These four models were retrospectively run from 1 January 1979 to 31 December 2008. Afterward, they are run in a near-real-time mode (with a three and half day lag).

The third stage moved toward evaluating and validating the quality and reliability of long-term NLDAS products using as many as available in situ observations and satellite-retrieved products. Evaluations span multiple spatial and temporal scales, from short term to long term, hourly to annual, and site and basin to continental United States. These observations include energy fluxes (e.g., downward shortwave and longwave radiation, upward shortwave and longwave radiation, net radiation, sensible heat flux, latent heat flux, ground heat flux, and such), water fluxes (e.g., evapotranspiration and streamflow), and state variables (e.g., soil moisture, soil temperature, land surface/skin temperature, snow water equivalent, and snow cover fraction). Recent works document progress toward evaluation and validation of NLDAS results. Comprehensive evaluation and comparison were detailed in *Xia et al*. [2012a, 2012b]. Overall results show that the products generated from NLDAS-2 have better quality when compared to those generated from NLDAS-1, due to both model and surface forcing data improvements. The simulated total runoff was evaluated against the observed streamflow at 986 small-medium (basin area $< 10{,}000\ km^2$) size basins and eight large size basins (basin area $\geq 10{,}000\ km^2$) which were measured by the U.S. Geological Survey (USGS). In the western coastal areas and in the eastern U.S., all four models are able to capture the broad features of observed streamflow. The four-model ensemble mean outperforms any individual model in terms of errors. A similar conclusion can be found for the validation of simulated evapotranspiration. The simulated soil moisture was evaluated using three observational data sets [*Xia et al*., 2014]: 20 year (1985–2004) monthly mean soil moisture from Illinois (17 sites), 6 year (1997–2003) daily mean soil moisture from the Oklahoma Mesonet (72 sites), and 8 year (2002–2009) daily soil moisture from the U.S. Department of Agriculture Soil Climate Analysis Network (121 sites). The results show that simulation skills of all four models are quite good in terms of anomaly correlation for both daily and monthly time scales, although the simulated soil moisture magnitude shows large errors, where some models may overestimate and other models may underestimate observed soil moisture. Similar to the streamflow and evapotranspiration evaluations, the four-model ensemble mean shows the most robust simulation skills over the continental United States when compared to any individual model.

The focus of the fourth stage is to apply long-term NLDAS products to support the National Integrated Drought Information System (http://www.drought.gov) and U.S. operational drought monitoring and prediction. One key application of the near-real-time NLDAS is in its drought monitoring over the continental United States, shown at the "NLDAS Monitor" tab of the NLDAS website (*Sheffield et al*., 2012; NCEP/EMC NLDAS website: http://www.emc.ncep.noaa.gov/mmb/nldas/; NASA NLDAS website: http://ldas.gsfc.nasa.gov/nldas). At the same time, the NLDAS team also routinely provides four-model ensemble mean daily, weekly, and monthly percentiles of the top 1 m soil moisture, total column soil moisture, total runoff, and evapotranspiration, which were widely used in drought analyses [*Andreadis et al*., 2005; *Shukla and Wood*, 2008; *Wang et al*., 2009] by the U.S. Drought Monitor (USDM) author group to help directly support the making of the weekly USDM. This team also provides NLDAS drought indices to support CPC monthly drought briefings and seasonal drought outlooks. However, these NLDAS drought indices are not comprehensively assessed as there are few reference drought data sets. The operational USDM (http://droughtmonitor.unl.edu/), an operational product [*Svoboda et al*., 2002; *Heim*, 2002], has generated many statistics (i.e., drought area percentages for the 48 states). How to best use and interpret these statistics, given the short period of record (2000 to present), in order to improve U.S. operational drought monitoring is still a challenging issue. This study will develop an objectively blended approach by establishing the linkage between NLDAS products and USDM statistics. The approach will use an optimization method to search for optimally blended weights and equations by minimizing the root-mean-square error (RMSE) between drought area percentage derived from NLDAS and from USDM. In turn, the USDM drought area percentage will be used to evaluate simulation skills of the optimal blended NLDAS drought index (OBNDI).

Table 1. The States That are Included in Each of the Six USDM Regions[a]

USDM Region	States (Acronyms) Included
West	Arizona (AZ), California (CA), Colorado (CO), Idaho (ID), Montana (MT), New Mexico (NM), Nevada (NV), Oregon (OR), Utah (UT),Washington (WA), Wyoming (WY)
High Plains	Colorado (CO), Kansas (KS), North Dakota (ND), Nebraska (NE), South Dakota (SD), Wyoming (WY)
South	Arkansas (AR), Louisiana (LA), Mississippi (MS), Oklahoma (OK), Tennessee (TN), Texas (TX)
Midwest	Iowa (IA), Illinois (IL), Indiana (IN), Kentucky (KY), Michigan (MI), Minnesota(MN), Missouri (MO), Ohio (OH), Wisconsin (WI)
Southeast	Alabama (AL), Florida (FL), Georgia (GA), North Carolina (NC), South Carolina (SC)
Northeast	Connecticut (CT), Delaware (DE), Massachusetts (MA), Maryland (MD), Maine (ME), New Jersey (NJ), New Hampshire (NH), New York (NY), Pennsylvania (PA), Rhode Island (RI), Virginia (VA), Vermont (VT), West Virginia (WV)

[a]Note that CO and WY are included in both the West and High Plains regions, as defined by the USDM.

The weekly USDM is a composite indicator, combining several variables into a single product that attempts to show both short- and long-term drought on one map. Variables (indices and indicators) utilized in the process address precipitation, temperature, vegetation health, soil moisture (modeled and in situ where available), streamflow, snowpack, snow water equivalent, reservoirs, and groundwater. The USDM is also unique in that it incorporates feedback and input into the process by maintaining and utilizing an expert user group of around 350 people in the field who serve as a "ground truth" to the product. A convergence of evidence approach is used to combine the scientific data with impacts and feedback from experts in the field via an iterative process. Strictly speaking, the USDM [*Svoboda et al*., 2002] is a state-of-the-art drought monitoring tool in the United States, which uses objective data sets combined with impacts and other subjective expert input in analyzing drought. Although the USDM captures drought well and is a powerful tool to monitor U.S. drought, it has a subjective element and cannot be reproduced in an objective way. In addition, there are no long-term USDM data as the USDM began in 2000. The main purpose of this work is to develop an automated USDM-based approach to objectively generate and reconstruct USDM-style maps using NLDAS-2 data sets by mimicking 10 year (2000–2009) USDM statistics.

This paper is organized as follows. The next section gives background on the data and methods used in this study. Section 3 describes the experimental design, optimal weights, and equations. Section 4 compares USDM statistics to those derived from OBNDI, examines the effect of Standardized Precipitation Index and snow water equivalent on the OBNDI, reconstructs long-term OBNDI, analyzes drought area percentages, and gives several drought examples. Section 5 gives a discussion about the effect of cold and warm season on the OBNDI. Section 6 summarizes this study and gives conclusions.

2. Data and Methods

2.1. Data

The data used in this study include NLDAS-2 monthly ensemble mean percentiles (http://www.emc.ncep.noaa.gov/mmb/nldas/drought/), USDM statistics, and NLDAS-2 mask data (i.e., land mask and state mask). For NLDAS-2 data, we used monthly mean percentiles of top 1 m soil moisture (SM1), total column soil moisture (SMT), total runoff (Q), and evapotranspiration (ET) as used by USDM author group. In addition, NLDAS-2 monthly snow water equivalent (SWE) percentile, and the 3 month and 6 month Standardized Precipitation Index (SPI) [*McKee et al*., 1993] values (hereafter, spi3 and spi6, respectively) are also used to examine their impact on the optimal blend by several sensitivity tests. A 13 year (2000–2012) USDM statistics (weekly mean drought area percentages) was downloaded from the USDM archive website (continental United States and 48 states: http://droughtmonitor.unl.edu/dmshps_archive.htm; six USDM regions: http://droughtmonitor.unl.edu/dmtabs_archive.htm). Each of the six USDM regions includes 5–13 states (see Table 1). In this study, we used number of days as weights (for the weeks that cross 2 months) to convert weekly mean drought area percentage to monthly mean drought area percentage. Five drought categories are classified as D0-D4 (percentile ≤ 30%), D1-D4 (percentile ≤ 20%), D2-D4 (percentile ≤ 10%), D3-D4 (percentile ≤ 5%), and D4 (percentile ≤ 2%). It should be noted that strictly speaking, only the last four categories can be classified as droughts; the first category is the abnormally dry (D0) case as defined by the USDM. For convenient statement and discussion, hereafter we use five drought categories (four drought categories and one

abnormally dry case) as used in *Anderson et al*. [2013] and *Xia et al*. [2013]. The NLDAS blended drought index is calculated as

$$\mathrm{OBNDI} = W_1\mathrm{SM1} + W_2\mathrm{SMT} + W_3\mathrm{Q} + W_4\mathrm{ET}, \tag{1}$$

where W_1, W_2, W_3, and W_4 are weight coefficients for four different variables separately, and SM1 is top 1 m soil moisture, SMT is total soil moisture, Q is runoff, and ET is evapotranspiration. The one-eighth-degree-gridded NLDAS-2 fields are geographically masked to calculate monthly drought area percentages for CONUS, six USDM regions, and 48 states for five drought categories. The weight coefficients are obtained by the optimization process described below. In separate sensitivity tests, we have also included spi3, spi6, and snow water equivalent (SWE) in the blended drought index to examine their effects (discussed in further detail later in this manuscript). The basic idea is to find optimal weight coefficients to minimize the error between drought area percentages derived from the USDM and those derived from the NLDAS blended drought index.

2.2. Optimization Approach

The optimization approach used in this study is the very fast simulated annealing (VFSA) algorithm. Details of VFSA have been described by many scientists [*Sen and Stoffa*, 1996; *Xia et al*., 2004], so only a brief description is given here. One may use the temperature constructed within the Metropolis algorithm [*Metropolis et al*., 1953] to locate the global minimum of an error function as defined in the following section by very slowly lowering the temperature parameter within

$$P = \exp\left(\frac{-\Delta E}{T}\right), \tag{2}$$

where P is the probability of acceptance of a new parameter set with positive change of error function values, ΔE is change of error functions calculated by new and previous parameter sets (see section 3.1), and T is a control parameter analogous to temperature. As used in a previous study [*Xia et al*., 2004], T is set as 3.0. If the change is negative, this new parameter set is accepted. If the change is positive, and if and only if P is less than a randomly generated number between 0 and 1, the new parameter set is rejected. This iterative process is analogous to the annealing process within a physical system where the lowest energy state between atoms or molecules is reached by the gradual cooling of the substance within a heat bath. Because of this physical analogy, the algorithm is called "Simulated Annealing". To enhance the ability of simulated annealing to converge to the global minimum of the error function, *Ingber* [1989] introduced a new procedure for selecting parameter sets according to a temperature-dependent Cauchy distribution. This modified simulated annealing algorithm is called very fast simulated annealing. The modified algorithm is described as follows.

Let us assume that model parameter m_i at kth iteration is represented by m_i^k such that

$$m_i^{\min} \le m_i^k \le m_i^{\max}, \tag{3}$$

where $m_i^{\min}$ and $m_i^{\max}$ are the minimum and maximum values of the model parameter m_i. This model parameter value is perturbed at iteration $k+1$ using $m_i^{k+1} = m_i^k + z_i\left(m_i^{\max} - m_i^{\min}\right)$, $m_i^{\min} \le m_i^{k+1} \le m_i^{\max}$ and $z_i \in [-1, 1]$. z_i is generated from the distribution

$$g_T(z) = \prod_{i=1}^{\mathrm{NM}} \frac{1}{2(|z_i| + T_i)\ln\left(1 + \frac{1}{T_i}\right)}, \tag{4}$$

and has a cumulative probability

$$G_{T_i} = \frac{1}{2} + \frac{\mathrm{sgn}(z_i)}{2}\frac{\ln\left(1 + \frac{|z_i|}{T_i}\right)}{\ln\left(1 + \frac{1}{T_i}\right)}, \tag{5}$$

where NM is the number of model parameter sets. *Ingber* [1989] showed that for such a distribution, the global minimum can be statistically obtained by using the cooling schedule

$$T_i(k) = T_{0i}\exp\left(-c_i k^{\frac{1}{\mathrm{NM}}}\right), \tag{6}$$

where T_{0i} is the initial temperature for model parameter i and c_i is a parameter to be used to control the

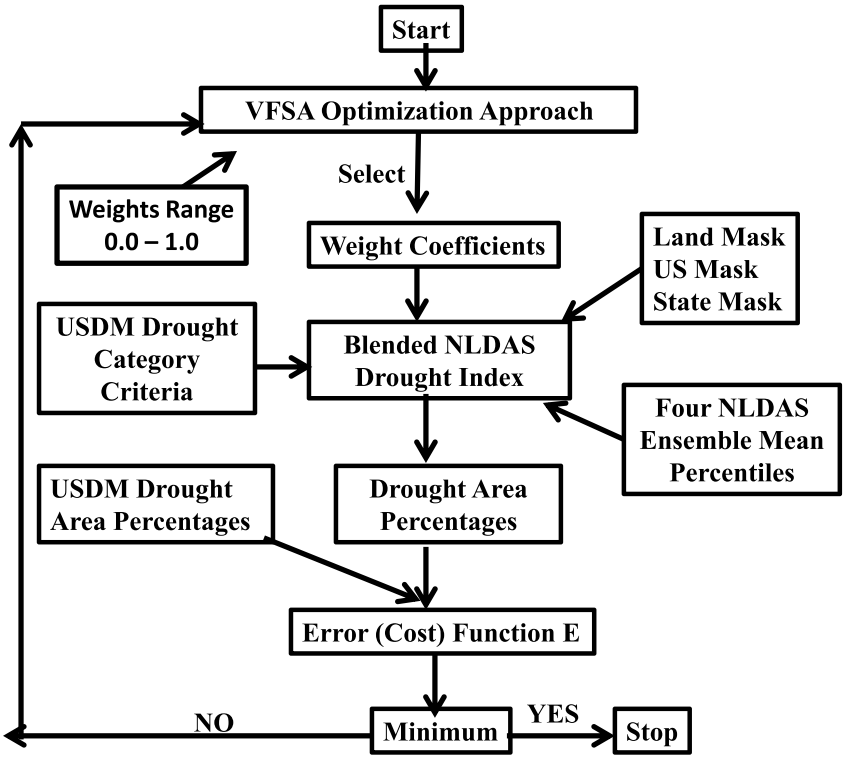

Figure 1. A schematic diagram to search for optimal weight coefficients by minimizing the error function between the drought area percentages derived from the USDM and blended NLDAS drought index.

temperature. The acceptance rule of the VFSA algorithm is the same as that used in the Metropolis rule. However, the VFSA is more efficient when compared with the simulated annealing algorithm.

2.3. Evaluation Metrics

The basic evaluation metrics for this study include bias, RMSE, correlation coefficients, and Nash-Sutcliffe efficiency [*Nash and Sutcliffe*, 1970]. The Nash-Sutcliffe efficiency (NSE) is defined as

$$\mathrm{NSE} = 1 - \frac{\sum_{t=1}^{\mathrm{MT}} (A_t - O_t)^2}{\sum_{t=1}^{\mathrm{MT}} (O_t - \overline{O})^2} \tag{7}$$

In equation (7) A_t and O_t are, respectively, drought area percentage derived from NLDAS and USDM, and $\overline{A}$ and $\overline{O}$ are their mean values for any given time period. The NSE is a measure of the drought area percentage simulation skill of the method as compared to the mean USDM drought area percentage and ranges in value from minus infinity (poor model skill) to one (perfect model skill). An efficiency of 0 (NSE = 0) indicates that the model simulations are as accurate as the mean of the USDM data, whereas an efficiency less than zero (NSE $<$ 0) occurs when the USDM mean is a better predictor than the model or, in other words, when the residual variance (described by the numerator in the expression above) is larger than the data variance (described by the denominator).

3. Optimization of Objectively Blended NLDAS Drought Indices

3.1. Experiment Design

In general, the error function (also called cost function) is defined as the RMSE between the observed and simulated data. However, observed drought data do not exist so far. The USDM [*Svoboda et al.*, 2002] is a state-of-the-art drought monitoring tool in the United States, which uses a combination of objective data sets and processes within a geographic information system environment along with impacts and input collected from over 350 experts in the field as input into a hybrid drought analysis. Therefore, monthly mean USDM drought area percentages are used as our reference data. Ten year (2000–2009) monthly drought area

Table 2. Optimal Weight Coefficients for CONUS, Region, and State Optimization Approach (Optimal Blended Drought Index = W_1SM1 + W_2SMT + W_3ET + W_4Q)[a]

U.S./Region	W_1	W_2	W_3	W_4	Cost
CONUS	**0.6253**	0.0253	0.0033	0.0000	0.0882
West	0.1083	**0.3935**	0.0000	0.0000	0.1674
High Plains	0.1940	**0.2816**	0.0000	0.0002	0.1380
South	0.2438	**0.3585**	0.0502	0.0000	0.0900
Midwest	**0.7551**	0.0757	0.0433	0.0175	0.0542
Southeast	0.1706	0.1490	0.0001	**0.3115**	0.1622
Northeast	**0.6651**	0.2571	0.0478	0.0027	0.0649

[a]Maximum weight coefficients are represented in bold.

percentages are used to construct our error function as a training period and those in the 2 year period from 2010 to 2011 are used for validation. The error function E can be defined as

$$E = 1/\mathrm{MT}\sum_{t=1}^{\mathrm{MT}}\sqrt{\frac{1}{C}\sum_{c=1}^{C}\left(A_{t,c}-O_{t,c}\right)^2}, \tag{8}$$

where MT is total number of months (120 in this study), C is the number of drought categories as described in section 2.1 (five in this study), and $A_{t,c}$ and $O_{t,c}$ are the drought area percentage derived from blended NLDAS drought index and USDM, respectively. The ranges of all four weights are selected from 0 to 1.

In this study, the three levels of USDM statistics data sets are utilized: CONUS, six USDM regions ("Region"), and 48 states ("State"). We optimized equation (1) for the continental United States, each of the six USDM regions, and each of the 48 states using an optimization process described in Figure 1. We also conducted cold and warm season sensitivity tests for nine states and eight sensitivity tests for spi3, spi6, and SWE. Therefore, there are 81 optimization tests in total. In order to achieve the optimization process convergence, each individual optimization test needs 1000 runs, and thus, we have 81,000 runs in this study. After the 1000 runs, if the VFSA does not convert to the required conditions, we select the weights with the minimum cost as the optimal weights. The purpose is to investigate if regional and/or state drought information can help enhance the accuracy of the OBNDI.

3.2. Optimal Weight Coefficients and Equations

The optimal weight coefficients and minimum error function values (costs) for CONUS and Region are shown in Table 2. The results show that SM1 plays a key role, SMT plays a small role, and Q and ET both play negligible roles in the continental United States. This has been discussed in recent work [*Xia et al*., 2013]. For the Region case, Q and ET play small or negligible roles in all six USDM regions, except for the Southeast where Q plays a key role. Either top 1 m soil moisture SM1 or total column soil moisture SMT or both play a key role for all six regions. Further analysis shows that SMT has larger weight coefficients than SM1 for the West, High Plains, and South. On the contrary, SMT has smaller weight coefficients than SM1 for the Midwest and Northeast. Over the Southeast, SMT and SM1 have similar magnitudes of weight coefficients, indicating they play the same role. We can see from the Region case analysis that different variables may play different roles in each of the six regions. This is a reasonable result as each of the six regions represents different climatic and geographical zones.

As discussed above, each of the six regions includes 5–13 states. Even within the same USDM region, the optimal weight coefficients are different when subregional/state climate, land cover, and soil property information are introduced into our optimization process. The results show that normalized optimal weight coefficients vary from state to state (Figure 2). For the State analysis, as is the case with the CONUS and Region analyses, ET plays a negligible role in all 48 states except for in WA, SD, LA, MS, OH, WV, NY, and MA where it plays an important role. Similar to ET, Q also plays a negligible role in all 48 states except for WY, MN, WI, MI, FL, SC, PA, RI, and DE, where Q plays an important and even dominant role. Although the VFSA is purely mathematical, the optimized results are relatively reasonable for Q as all nine states except for WY are nearby a coast or lake where streamflow is a large part of the annual total water balance. Again, either SM1 or SMT or both play the dominant role for all 48 states although one is more important than the other except for the six states with similar weights (i.e., WA, MT, CA, AR, GA, and MD). The states having the largest weight coefficients for SM1 are covered by crop land, which is consistent with its 1 m root zone parameter. The states which have

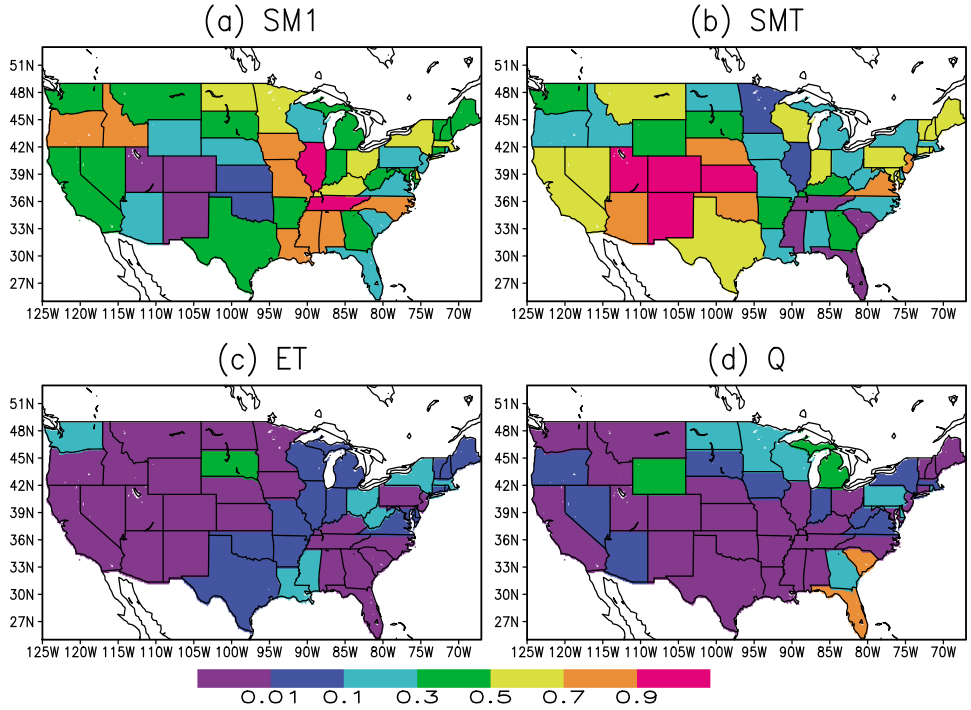

Figure 2. Spatial distribution of normalized optimal weight coefficients for (a) SM1, (b) SMT, (c) ET, and (d) *Q*.

the largest weight coefficients for SMT are mainly covered by shrubland and woody grassland. This is also consistent with their 1.5–2.0 m root zone, which is the depth of the total column soil moisture. Exceptions to this are found in three states, Oklahoma, Kansas, and Nebraska, which have large weight coefficients for SMT. Because these states have large agricultural and grassland regions, one might expect that SM1, rather than SMT, would play a dominant role. The reason for this exception remains unclear and needs to be studied further. A reasonable hypothesis for this is that the current NLDAS system does not consider deep soil moisture recharge due to a lack of an interactive groundwater module in the LSMs and that NLDAS also does not include the effects of irrigation within the LSMs. Both of these effects which can be important in many locations, but at times especially in these states, may result in larger weight coefficients for SMT over SM1 in representing drought variability.

For the six USDM regions, we tested the addition of precipitation-related drought indices (e.g., spi3 and spi6) to the OBNDI, as spi3 and spi6 have been widely applied to drought analysis and monitoring [*Heim*, 2002; *Mo*, 2008; *Hayes et al*., 2011]. In addition, we also tested the addition of monthly SWE percentile values to the blended drought index experiment as recent work from the University of Washington (B. Nijssen et al., A prototype Global Drought Information System based on multiple land surface models, submitted to *Journal of Hydrometeorology*, 2013) has demonstrated that SWE has a significant impact on winter drought intensity

Table 3. Optimal Blended Drought Indices (OBNDI) are Selected for Six USDM Regions When SM1, SMT, spi3, spi6, and SWE are Considered Separately (OBNDI = W_1SM1 + W_2SMT + W_3spi3 + W_4spi6 for Six USDM Regions and OBNDI = W_1SM1 + W_2SMT + W_3ET + W_4SWE for USDM West Region for Winter Months)

Region	W_1	W_2	W_3	W_4	Cost
West	0.0049	0.4076	0.0030	0.0000	0.1695
High Plains	0.1089	0.1177	0.0001	0.0029	0.0885
South	0.0742	0.1595	0.0004	0.0042	0.0968
Midwest	0.5121	0.0980	0.0005	0.0024	0.0529
Southeast	0.5265	0.0115	0.0000	0.0018	0.1494
Northeast	0.6995	0.1580	0.0003	0.0020	0.0650
Winter months (December, January, February, and March)					
West	0.0956	0.0744	0.0171	0.4583	0.1771

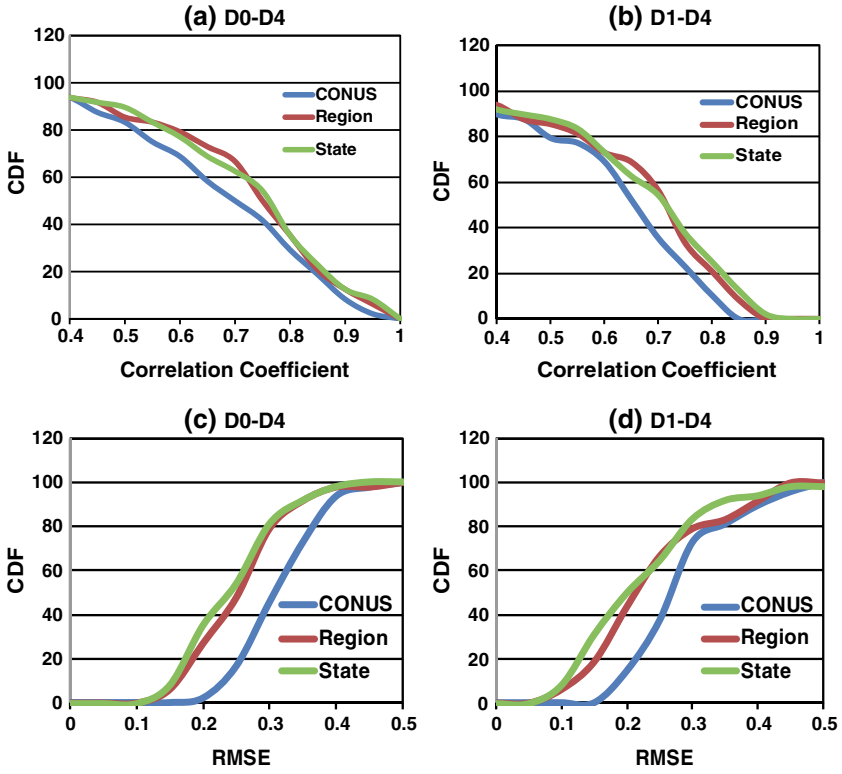

Figure 3. The cumulative density function (CDF) of correlation and RMSE between drought area percentages derived from the USDM and OBNDI for (a) correlation for D0-D4, (b) correlation for D1-D4, (c) RMSE for D0-D4, and (d) RMSE for D1-D4. The CDF is derived from 48 states for the training period 2000–2009.

in eastern Washington State. These sensitivity test results in Table 3 show that spi3, spi6, and SWE all have small weights ($<$ 5% of total weight) when compared to soil moisture for all six regions and to total runoff for the Southeast region. Precipitation has small weights because NLDAS soil moisture has a closer correlation with the USDM than precipitation, indicating the value of land surface models [*Anderson et al*., 2013]. The drought signal in precipitation has been incorporated into the modeled soil moisture. SWE has small weights because there is little SWE in South and Southeast; for the Northeast, Midwest, High Plains, and West, SWE plays a small role when compared with soil moisture in the context of an annual period, which is used for the optimization process. When spi3 or spi6 is blended into the OBNDI, the cost is reduced by 7.9% and 2.4% for the Southeast and Midwest regions, respectively. In contrast, for the High Plains and South regions, the cost is increased by 9.7% and 6.0%, respectively. For the West and Southeast region, there is little change. Further sensitivity testing shows that for the winter months (December, January, February, and March) over the western mountainous regions, SWE has the largest weight (71% of total weight), soil moisture has the second largest weight (27% of total weight), and ET has the smallest weight (2% of total weight). Unlike in the West, SWE during the winter over the other five regions plays a small role and carries a very small weight.

4. Comparison and Analysis

4.1. Comparison of OBNDI and USDM

The analysis in section 3.2 shows that optimal weight coefficients depend on region and state which have different climates and soil and vegetation types. This section will compare different optimal blended drought indices from CONUS, Region, and State spatial scales with USDM drought area percentages for the training period (2000–2009). Figure 3 shows the cumulative density function (CDF) of correlation (R) and root-mean-square error (RMSE) between NLDAS and USDM drought area percentage for D0-D4 and D1-D4 category. The results show that both Region and State have larger correlations (Figures 3a and 3b) and smaller RMSE values (Figures 3c and 3d) when compared to CONUS although State marginally outperforms Region for both D0-D4 and D1-D4 cases. Table 4 shows the number of states which have the smallest biases between USDM and OBNDI drought area

Table 4. Number of States With the Smallest Bias (Within a 0.1 Range) Between USDM and OBNDI Drought Percentage is Listed

Method	D0-D4	D1-D4	D2-D4	D3-D4	D4
CONUS	22	38	39	45	48
Region	41	42	40	44	48
State	45	46	38	46	48

percentage (within a 0.1 range) for five drought categories. The results show that State and Region have more states, which have a smaller bias when compared to CONUS, in particular for D0-D4 and D1-D4. When we compare total drought area percentages over the continental United States, State and Region have smaller RMSE values and larger correlations than CONUS for D0-D4 and D1-D4 and comparable RMSE and correlation for D2-D4 and D3-D4, although State has worse performance than CONUS (Figure 4). These analyses show that overall, State and Region have better performance than CONUS as they have larger correlation values and smaller RMSE and bias. The physical reason State and Region have better performance than CONUS is that the former incorporates more drought indices and spatially varied weighting (depending on USDM region and state, Table 2 and Figure 2) into the OBNDI whereas the latter only includes the top 1 m soil moisture and constant weight without considering climate (e.g., spatial distribution of precipitation), land cover (e.g., crops, grasslands, woodland, and forest), and soil type (e.g., sand, clay, and loam) within the different USDM regions and states (Table 2). The weights may vary from region to region as vegetation type varies. For example, agricultural drought may vary from region to region and season to season as a plant's demand for water is

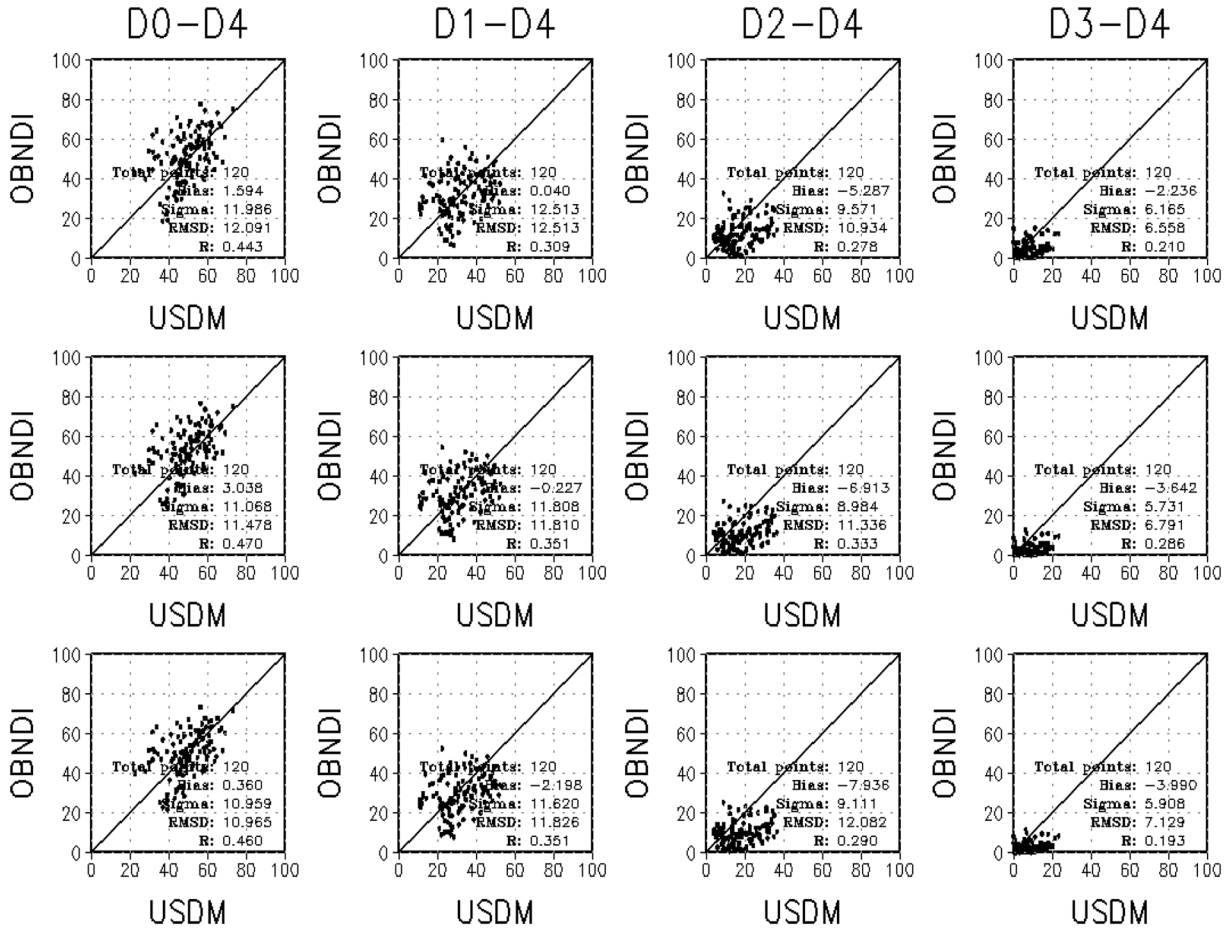

Figure 4. Comparison between the drought area percentages over the continental United States derived from the USDM and OBNDI during the training period 2000–2009. (top) CONUS. (middle) Region. (bottom) State. The number in the bottom right of each plot shows total months, bias, root-mean-square error (RMSE), standard deviation (sigma), and correlation (*R*).

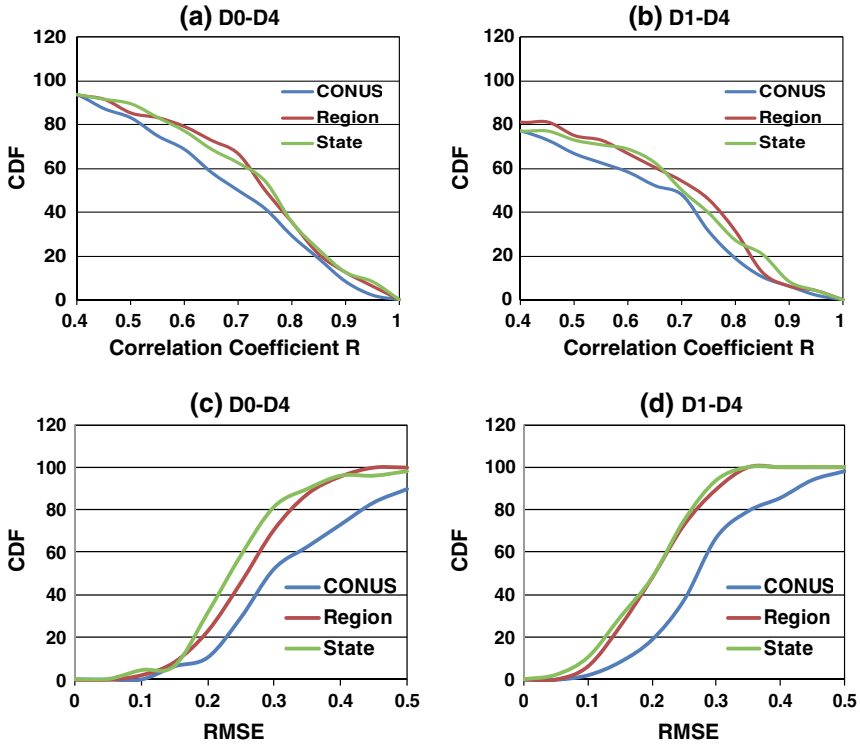

Figure 5. The same as Figure 3 except for the validation period 2010–2011.

dependent on prevailing weather conditions, biological characteristics of the specific plant, its stage of growth, and the physical and biological properties of the soil.

4.2. Validation of OBNDI

For the training period, the results show that State and Region perform better than CONUS and State marginally outperforms Region. This section will investigate if this conclusion is true for an independent validation period (2010–2011). Note that during this validation period, the NLDAS data sets were sent to the USDM authors, so separating this period keeps the OBNDI truly independent of the USDM for these 2 years. Figure 5 shows the CDF of *R* and RMSE between NLDAS and USDM drought area percentage for D0-D4 and D1-D4 categories. By comparing Figure 5 with Figure 3, the results are relatively similar, although there are larger correlations and RMSE. Table 5 shows the overall statistics (i.e., *R*, bias, and RMSE) of drought area percentage over the continental United States for CONUS, Region, and State for the validation period. State

Table 5. Statistics of CONUS Drought Area Percentage Derived From the USDM and Three Methods for Five Drought Categories During the Validation Period (2010–2011)[a]

Statistics	D0-D4	D1-D4	D2-D4	D3-D4	D4
CONUS					
Bias (%)	14.4	12.6	4.0	1.2	0.3
RMSE (%)	16.2	14.3	6.3	3.0	2.2
R	0.60	0.75	0.89	0.89	0.89
Region					
Bias (%)	14.8	11.7	2.8	−0.1	−0.8
RMSE (%)	16.5	13.2	5.0	3.5	2.7
R	0.61	0.79	0.93	0.94	0.88
State					
Bias (%)	12.8	10.5	2.0	−0.5	−0.7
RMSE (%)	14.7	12.0	4.3	3.5	2.5
R	0.63	0.80	0.95	0.95	0.90

[a]In this study 24 monthly averaged drought area data are used.

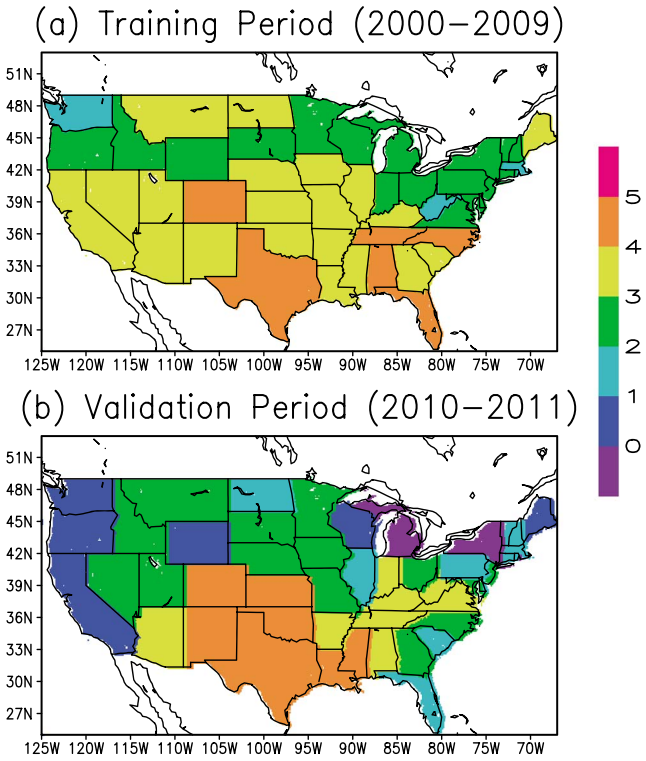

Figure 6. Spatial distribution of number of categories with significant correlation at the 95% confidence level for (a) training period 2000–2009 and (b) validation period 2010–2011. The correlation threshold value is 0.2 for the training period and 0.4 for validation period. The result is for the State case.

and Region have larger correlations and smaller errors when compared to CONUS. State has the best performance, followed by Region and CONUS, which is consistent with that derived from the training period.

4.3. Analysis of Several Drought Examples

Figure 6 shows the spatial distribution of the number of categories with significant correlations at the 95% confidence level for the training period (Figure 6a) and validation period (Figure 6b). The number 5 represents a significant correlation for all five categories, a number 4 represents a significant correlation for four categories, and so on. The results show that for the training period, the states in the South and Southeast regions have the most significant correlation, while the middle- and high-latitude states have the least correlation. For the validation period, the states with the most significant correlation appear in the South and Southeast regions although the states may differ for the training and validation period. Four states do not show a significant correlation between the USDM and NLDAS drought area percentage for all five categories, and six states have significant correlations for only one drought category. The largest deterioration of drought percentage area correlation occurs in FL, CA, MI, and NY. The other poorly performing states are WA, OR, WY, WI, ME, and RI. This means that the OBNDI may not be stable for these states as some of them are streamflow dominated (i.e., FL, MI, WI, and WY), and thus, a longer (e.g., 30 years) USDM drought area percentage data set may be needed for determining a stable OBNDI. In addition, the poor performance of OBNDI in western states may be associated with poor NLDAS simulations due to "inaccurate" gauge precipitation (e.g., significant reduction of gauge stations post-2002 [*Mo et al.*, 2012]) and effects of complicated topography and snowpack processes. This will need further investigation.

Figure 7 shows the comparison between the USDM and NLDAS drought area percentages (D1-D4) for nine states. These states have the highest correlation for both training and validation periods, and the results are very encouraging. The OBNDI can capture variability of monthly drought events very well. Figure 8 shows the comparison in the same nine states for the validation period. Although NLDAS overestimates drought severity in NE and KS, it captures the variability and magnitude of monthly mean drought area percentage

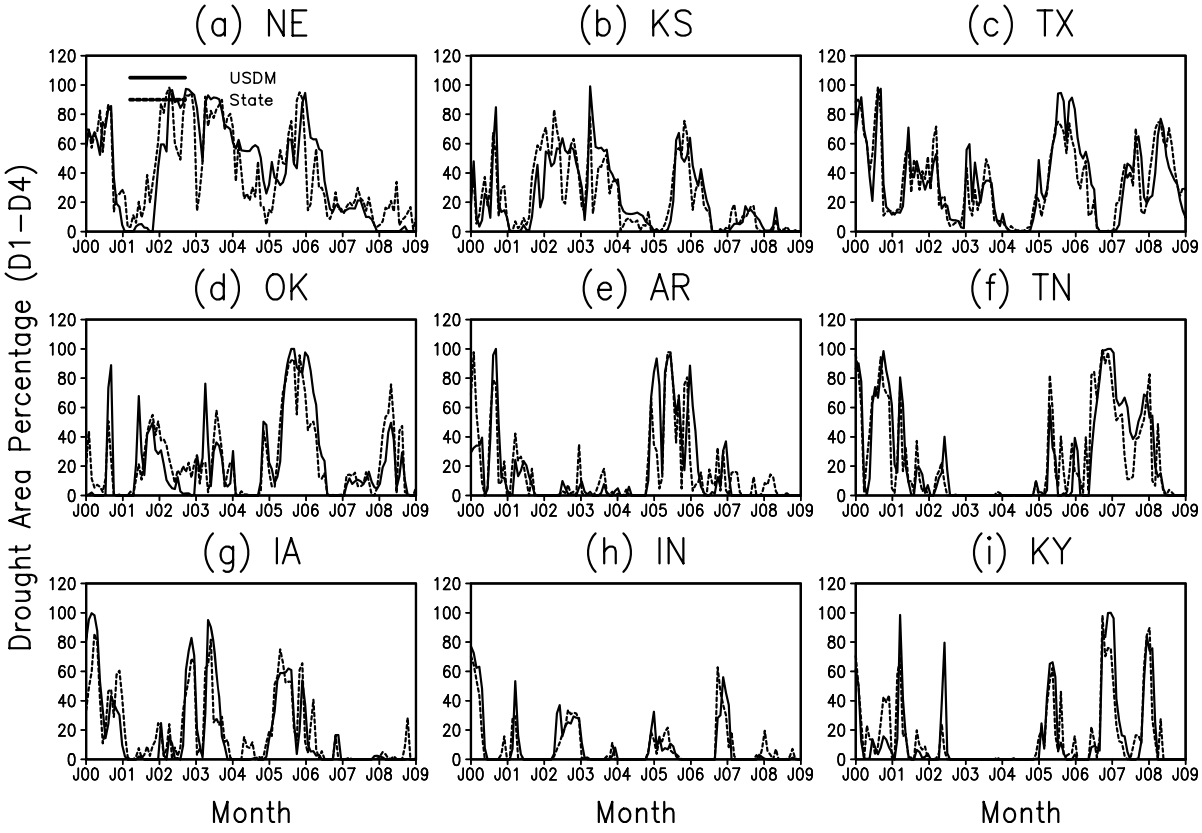

Figure 7. Comparison between drought area percentages (D1-D4) derived from USDM (solid line) and OBNDI (State, dashed line) for the training period 2000–2009 for the nine states with the highest correlations for both the training and validation periods.

well for the other seven states. A further comparison between the USDM and NLDAS for all five drought categories and TX is shown in Figure 9. The results show that NLDAS captures broad features and variability of drought area percentages well (Figure 9b) when compared to USDM drought area percentages (Figure 9a).

Figure 10 gives the spatial distribution of Nash-Sutcliffe efficiency in 48 states for the training (Figures 10a, 10b, and 10c) and the validation periods (Figures 10d, 10e, and 10f). When NSE is larger than 0.4, as suggested by *Xia et al.* [2012b], this means that the OBNDI has useful skills in simulating USDM drought area percentages for that state and that drought category. If the NSE is larger than 0.0 but smaller than 0.4 (in green), this means that the drought area percentage derived from the OBNDI is a better simulator than the mean drought area percentage derived from the USDM. If NSE is smaller than 0.0, this means that there is no skill from the OBNDI and a 10 year mean USDM drought area percentage should be used for that state and drought category. The results show that for the training period, most of the 48 states have useful skills for D0-D4 and D1-D4 and many of the 48 states have useful skills for D2-D4. There are only a few states which have no apparent skill, such as WA, WY, ID, UT, AZ, SD, and WV. In contrast, for the validation period, most of the 48 states, except for those in the South, have no skill for D1-D4 and D2-D4 although many of the 48 states have skill for D0-D4. The main reason is that the optimization method minimizes the RMSE of the drought area percentage derived from blended NLDAS drought index and the USDM for the training period. However, there are much larger RMSE for the validation period than the training period as shown in Figures 7 and 8, although the correlation coefficients are similar for the two periods (Tables 3 and 4). For the other two drought categories (i.e., D3-D4 and D4), only a few have useful simulation skills for the training period, for instance ND (0.431), KS (0.456), MO (0.505) for D3-D4, and MN (0.667) for D4. For the validation period, five states can simultaneously capture variability and magnitude of the drought area percentage well. The NSE for D3-D4 is 0.623, 0.842, 0.408, 0.426, and 0.714 for NM, TX, OK, WV, and PA, respectively. The value for D4 is 0.5, 0.661, 0.312, 1.0, and 1.0 in these same states. Overall, the OBNDI has difficulty in capturing the variability and magnitude of extremely severe drought events for almost all states. The major reason for this is that there is a very limited sample size for D2-D4, D3-D4, and D4 categories as drought area percentage is either zero or a

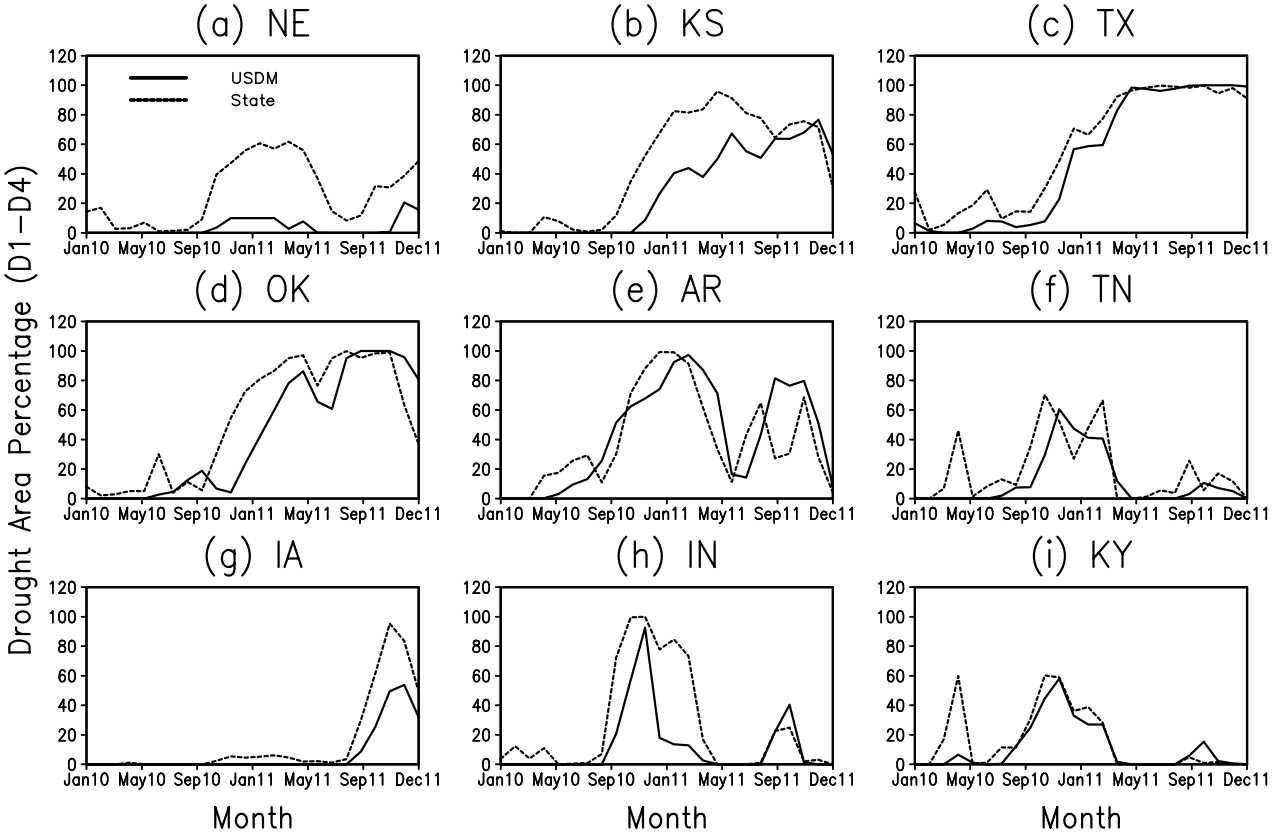

Figure 8. The same as Figure 7 except for the validation period 2010–2011.

very small value for these categories. Therefore, more effort needs to be invested into enhancing the capacity of the OBNDI by adding independent drought indicators such as observations (i.e., observed streamflow), remote sensing (e.g., Evaporative Stress Index [*Anderson et al.*, 2011], Vegetation Drought Response Index [*Brown et al.*, 2008], and Gravity Recovery and Climate Experiment-based groundwater storage [*Houborg*

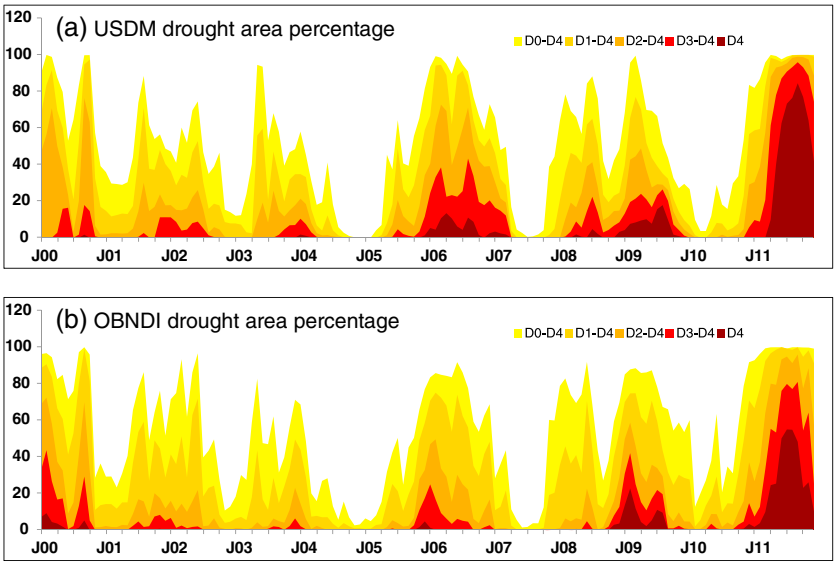

Figure 9. Comparison of drought area percentages of (a) USDM and (b) State for five drought categories in Texas from January 2000 to December 2011.

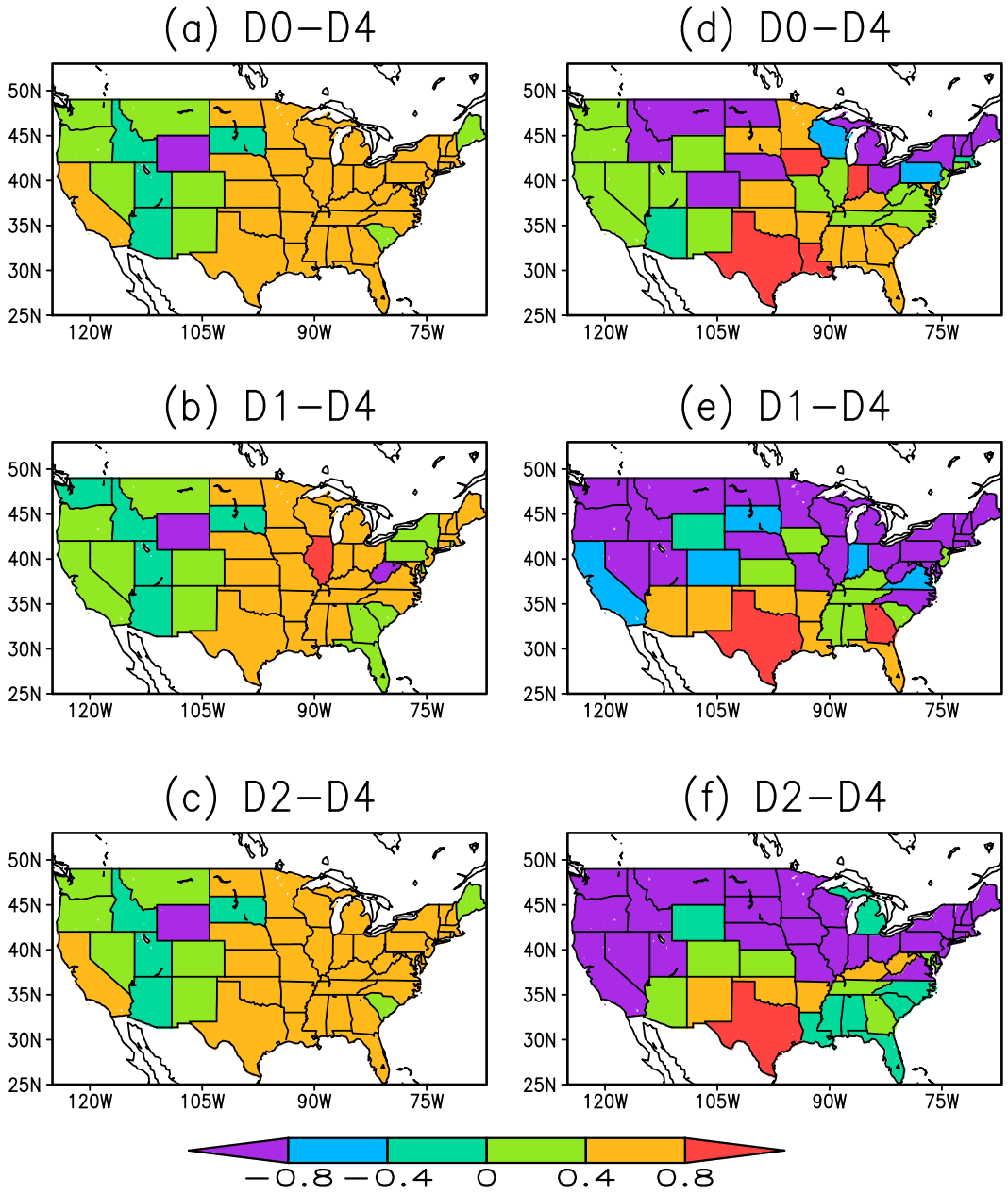

Figure 10. Nash-Sutcliffe efficiency for drought categories: (a) D0-D4, (b) D1-D4, (c) D2-D4 during the training period (2000–2009), (d) D0-D4, (e) D1-D4, and (f) D2-D4 during the validation period (2010–2011).

et al., 2012]), and other drought indicators (e.g., Standardized Precipitation-Evapotranspiration Index [*Vicente-Serrano et al*., 2010], Palmer Drought Severity Index [*Karl*, 1986], Palmer Modified Drought Index [*Mo and Chelliah*, 2006], and Palmer Hydrological Drought Index [*Heim*, 2002]).

4.4. Reconstruction of 33 Year (1980–2012) Drought Area Percentage Over the Continental United States

Based on the optimal weight coefficients and equations derived from each of the 48 states, we can reconstruct long-term drought area percentages for five drought categories within each state. However, some states may have more reliable/stable optimal equations and more useful simulation skills than the others as discussed in section 4.5. Still, theoretically this reconstruction can be completed using monthly ensemble mean NLDAS drought indices by assuming that weight coefficients do not change with the

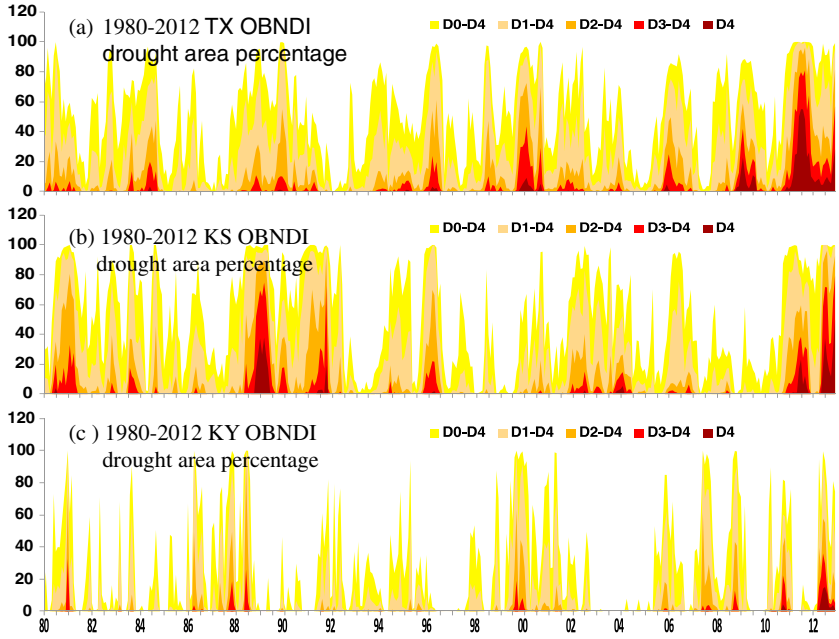

Figure 11. Thirty-three year (1980–2012) reconstructed drought area percentages for five categories in Texas, Kansas, and Kentucky. These drought area percentages are derived from the State experiment.

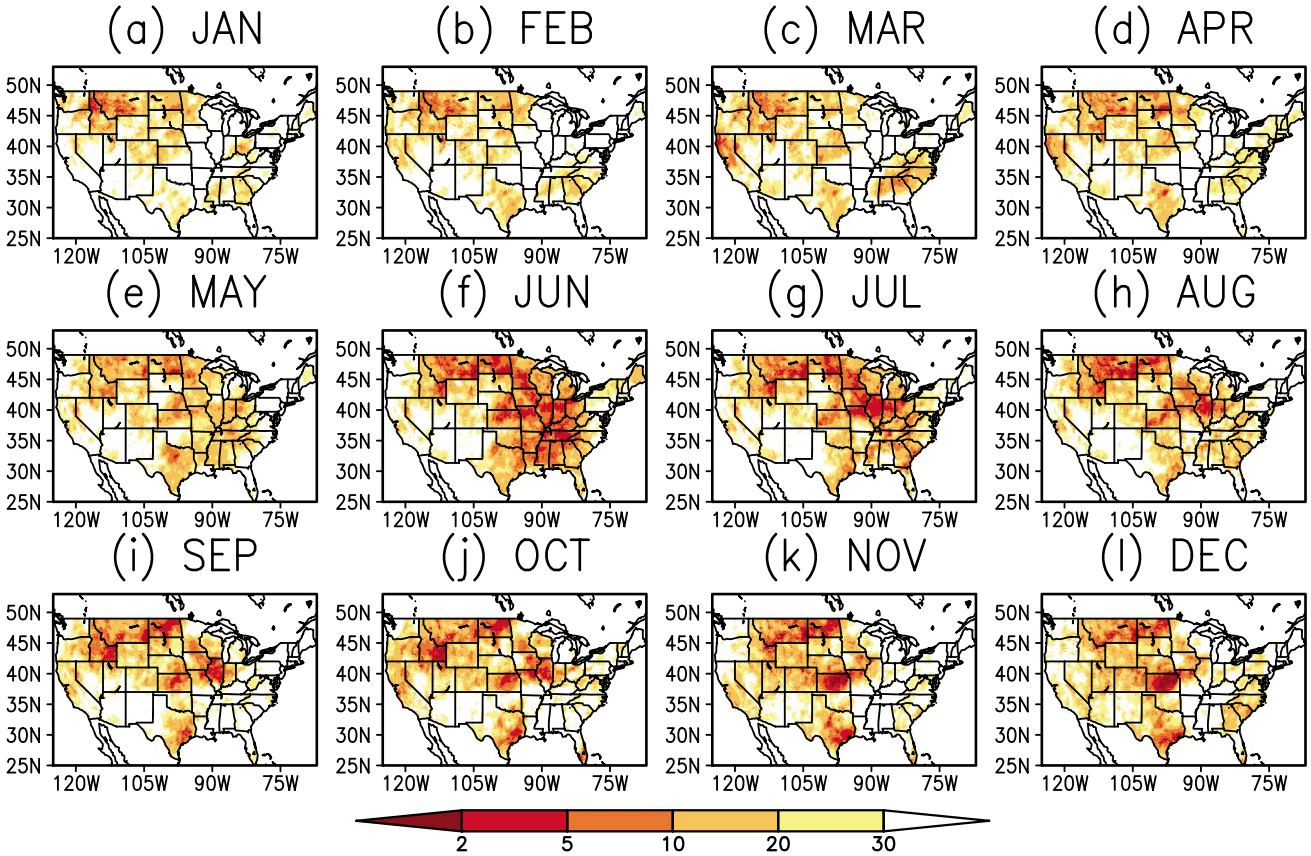

Figure 12. January to December 1988 OBNDI derived from the State experiment.

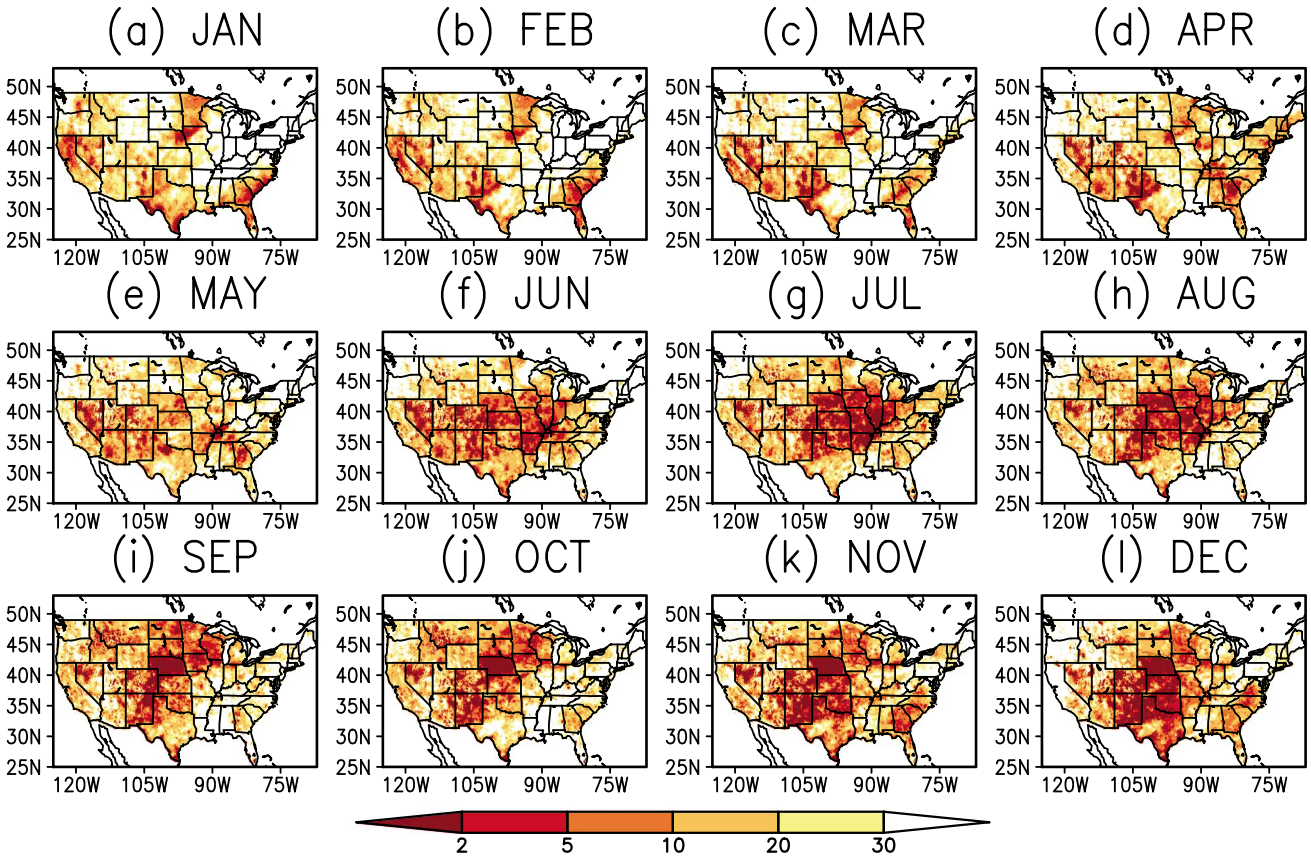

Figure 13. January to December 2012 OBNDI derived from the State experiment.

decadal climate change. It should be noted that this assumption may lead to the errors of the reconstructed drought area percentage and OBNDI. However, the reconstructed OBNDI is the best proxy based on our current knowledge. The reconstructed drought area percentages for three states, TX, KS, and KY, are shown in Figure 11. These three states represent the South, High Plains, and Midwest regions, respectively. The results show that the most severe droughts occurred in 2000, 2006–2007, 2009, and 2011 for TX, in 1989, 2011, and 2012 for KS, and in 2011 and 2012 for KY. Incorporating the same process, we can reconstruct the drought area percentage for all 48 states and this reconstruction process is reproducible.

4.5. Analysis of 33 Year (1980–2012) Reconstructed Gridded OBNDI

Thirty-three year reconstructed OBNDI drought area percentage has been analyzed in section 4.4. Three examples have shown that, in general, OBNDI can reasonably capture monthly variability and magnitude of drought area percentage for three states. This section will examine if OBNDI can capture spatial pattern of drought events using 33 year reconstructed gridded OBNDI. We analyzed all 33 year gridded OBNDI and found that it can broadly capture the spatial patterns of drought events and their intensification, expansion, and termination. The 1988 and 2012 drought events are used as two examples below.

4.5.1. The 1988 Drought

As an example, Figure 12 shows the variation of the 12 month OBNDI in 1988. From January to April, the drought occurred in the High Plains, Northwest, and eastern Texas and there is little variability from month to month for areal coverage and intensity. From the beginning of May, the drought suddenly increased and quickly extended to the Midwest and Ohio Valley. In June and July, drought severity further intensified and the area coverage expanded. From the beginning of August, the drought severity and area were reduced. From September through December, the drought's severity and area were virtually unchanged, keeping the

Table 6. Weight Coefficients and Cost Values for Nine States and Two Seasons (Cold and Warm)

State	W_1 for SM1 Warm (Cold)	W_2 for SM2 Warm (Cold)	W_3 for ET Warm (Cold)	W_4 for Q Warm (Cold)	Cost Warm (Cold)
CO	0.0823(0.1242)	0.3694(0.3949)	0.0001(0.0003)	0.0000(0.0031)	0.2208(0.2150)
KS	0.0309(0.0614)	0.5087(0.4343)	0.0000(0.0024)	0.0001(0.0175)	0.1000(0.0914)
OK	0.5421(0.0918)	0.1821(0.5749)	0.0001(0.0523)	0.0002(0.0033)	0.1129(0.1059)
IL	0.7508(0.5809)	0.1545(0.4135)	0.0080(0.0115)	0.0002(0.0001)	0.0771(0.0766)
TX	0.5083(0.1200)	0.0604(0.4028)	0.0001(0.0484)	0.0011(0.0128)	0.1188(0.1047)
IA	0.5457(0.8424)	0.3386(0.0784)	0.0010(0.0072)	0.0532(0.0485)	0.0762(0.0987)
MT	0.0928(0.3716)	0.2908(0.0188)	0.0000(0.0000)	0.0001(0.0001)	0.2174(0.2230)
NH	0.1078(0.3821)	0.3876(0.5680)	0.0012(0.0356)	0.4027(0.0045)	0.0802(0.1005)
FL	0.1990(0.0617)	0.2835(0.4028)	0.0001(0.0062)	0.1470(0.8054)	0.1894(0.1617)

drought in the Northwest, High Plains, Midwest, and Southern Texas. This suggests that the OBNDI is able to capture the broad features and monthly variations of the 1988 drought's severity and area extent.

4.5.2. The 2012 Drought

The 2012 United States drought is an expansion of the 2010–2012 Southern United States drought, which began in the spring of 2012 when the record-shattering (the lowest SWE) lack of snow in the United States led to very little meltwater being absorbed into the soil. The USDM results show that it peaked at 80% of the continental United States in July with at least abnormally dry (D0) conditions. Out of that 80%, 62% is designated as at least moderate drought (D1) conditions (the actual D1-D4 peak occurred during September at 65%). The 2012 drought affected a similarly large area as those droughts experienced in the 1930s and 1950s but was different in that it had not yet been in place as long. The drought exceeded, by most measures (e.g., drought severity and drought area coverage), the 1988–1989 case—the most recent comparable drought in the country's midsection. Figure 13 shows the drought's start and expansion in 2012 as depicted by the OBNDI. The results show that from January through March, the drought occurred in the West, parts of the Midwest, and the Southeast. From April to the end of the year, drought intensity and area rapidly expanded to two thirds of the country. The drought covered much of the West (except for the Pacific Northwest), the Great Plains, and parts of the Midwest and Southeast. Drought is the most severe in July and August as shown in Figure 13. The results are in good agreement with the U.S. Drought Monitor.

5. Discussion

Although this method shows a promising capacity to broadly simulate USDM drought area percentages, it still has several issues that need to be explored. In order to investigate the dependence of weight coefficients on season, we randomly selected nine states including CO, FL, IA, IL, KS, MT, NH, OK, and TX from which we separated 10 year (from 2000 to 2009) drought area percentage data into two seasons: a cold season (from October to March) and a warm season (from April to September). The weight coefficients and cost values for nine states and two seasons are listed in Table 6. In general, for both warm and cold seasons, there are small changes for cost values. For weight coefficients, three cases can be found. The first case is that there is little change in the weight coefficients (i.e., CO and KS) for both the warm and cold seasons, the second case reflects a redistribution of weight coefficients between SM1 and SMT for warm and cold seasons (i.e., OK, IA, IL, MT, and TX) when the sum of the weight coefficients for SM1 and SMT is similar, and the third case is that there are completely different weight coefficients for the warm and cold seasons (i.e., FL and NH). The latter two cases show a strong dependence of weight coefficients on season. This analysis demonstrates that the weight coefficients are seasonally dependent for most of states.

It should be noted that the results for the administrative area "states" are very useful for state-based management. However, as drought is a natural phenomenon including meteorological (lack of precipitation), agricultural (soil moisture deficit), and hydrological (shortage of streamflow and groundwater supplies) drought, it is closely related to climate and hydrological conditions for a specific region. Therefore, experiments conducted over hydroclimatic regions (e.g., U.S. climate divisions) may be more reasonable. Such a study will be conducted in future.

6. Summary and Conclusions

We extended our recent work [*Xia et al.*, 2013] from the continental United Sates to six USDM regions (i.e., West, South, High Plains, Midwest, Southeast, and Northeast) and 48 states. The results show that the

State and Region cases outperform the CONUS in terms of bias, RMSE, and correlation. The State marginally outperforms the Region experiment in terms of the same criteria utilized for both the training and validation periods. The optimal weights for different variables depend on the chosen region and/or state. Soil moisture plays a dominant role in total weight for most of the six USDM regions and 48 states. Total runoff plays an important role only in the Southeast region and nine other states (WY, MN, WI, MI, FL, SC, PA, RI, and DE), most of them are close to the coasts or the lakes where streamflow is a large part of annual water budgets. Evapotranspiration plays an important role in total weight within eight states (WA, SD, LA, MS, OH, WV, NY, and MA). The evaluation of NLDAS products has shown that the simulated total runoff is poor in most parts of United States, except for the states along the western coast and Southeast, and the simulated evapotranspiration is poor in winter and in forested regions when compared to the observations [*Xia et al*., 2012b]. However, the simulated soil moisture simulation is quite reasonable and reliable when compared to the observations (Xia et al., 2014). Therefore, a larger weight for soil moisture is a favorable result and thus may give more accurate drought indices.

Overall, our evaluation of the OBNDI derived from the State experiment shows that for the training period, there are significant correlations between drought area percentages (i.e., D0-D4 and D1-D4 categories) derived from the USDM and NLDAS for all 48 states. Further analysis shows that there are significant correlations for most states for the D2-D4 category. For the validation period, there are significant correlations for most states between the D0-D4 and D1-D4 categories. In terms of NSE, for the training period, large simulation skills appear in the states of the South, Southeast, High Plains, and Midwest region for the D0-D4, D1-D4, and D2-D4 categories. There are few simulation skills in the states of western and northeastern region for all five categories. The reason for the West region may be due to poor precipitation forcing as there are fewer gauge stations available after 2002 in the real-time period [*Mo et al*., 2012], which can lead to inaccurate simulations from land surface/hydrological models. Another possible reason is the effect of topography and snowpack-related processes. The reason for lack of skill in the Northeast region remains unclear. There are no skills for D3-D4 and D4 categories in almost all 48 states except for ND, KS, MO, and MN. For the validation period, there are a few states (i.e., NM, TX, OK, WV, and PA) with useful simulation skills (i.e., NSE $>$ 0.4) [*Xia et al*., 2012b] when compared to the training period for all categories. The reason for poor simulations during the validation period is that there are larger simulation errors when compared to the training period. The only exception is Texas, where there are large NSE values for all five categories (D0-D4: 0.84, D1-D4: 0.92, D2-D4: 0.95, D3-D4: 0.84, and D4:0.66) because of intense 2010–2011 drought events.

The analysis of reconstructed drought area percentages shows that the OBNDI is able to capture broad features of drought area percentages such as magnitude and monthly variability for the D0-D4 and D1-D4 in many states, which are mainly located in the South, Southeast, High Plains and Midwest regions. The reconstructed OBNDI can basically capture the start, duration, and termination of drought events. However, there is still significant room for improving and enhancing our simulation skills, in particular for most states of the western and northeastern regions, and for the most severe drought events (D3-D4 and D4). The impact of accurate gauge precipitation on simulation skills in western region will need to be addressed in the future by rerunning the four NLDAS models using a retrospective gauge precipitation data set. Irrigation and groundwater processes need to be added to the NLDAS framework to improve NLDAS simulations. In addition, after more independent inputs such as observed streamflow (e.g., percentiles) from USGS, remote-sensing drought indices, and operational drought indices will be used to blend with the four NLDAS drought indices used in this study, further improvement can be expected. Moreover, as suggested in the USDM experimental objective blends of drought indicators housed at the CPC (http://www.cpc.ncep.noaa.gov/products/predictions/tools/edb/droughtblends.php), drought can be classified as long term or short term according to different drought indicators. In order to improve the skills of the OBNDI, understanding the frequency of long-term and short-term drought occurrence for each state/USDM region, and for each season, is also an important step because the understanding can help in the selection of the most appropriate drought indicators/triggers. It should be noted that the OBNDI is easily reproducible, which is quite different from the USDM, which is based on a combination of objective and subjective analyses, making it difficult to reproduce. Therefore, this approach can be used to reasonably reconstruct long-term drought area percentages and derivative drought indices.

Acknowledgments
The NLDAS project is sponsored by the Modeling, Analysis, Predictions, and Projections (MAPP) Program within NOAA's Climate Program Office. The authors thank Weizheng Zhen, Helin Wei, and three anonymous reviewers whose edits and comments greatly improved the quality and readability of this manuscript. Y.X. also thanks Kingtse Mo from Climate Prediction Center who helped compute spi3 and spi6.

References

Anderson, M. C., et al. (2013), An intercomparison of drought indicators based on thermal remote sensing and NLDAS-2 simulations with U.S. Drought Monitor classifications, *J. Hydrometeorol.*, *14*, 1035–1056.

Anderson, M. C., C. Hain, B. Wardlow, A. Pimstein, J. Mecikalski, and W. P. Kustas (2011), Evaluation of a drought index based on thermal remote sensing of evapotranspiration over the continental U.S., *J. Clim.*, *24*, 2025–2044.

Andreadis, K. M., E. A. Clark, A. W. Wood, A. F. Hamlet, and D. P. Lettenmaier (2005), Twentieth-century drought in the conterminous United States, *J. Hydrometeorol.*, *6*, 985–1001.

Brown, J. F., B. D. Wardlow, T. Tadesse, M. J. Hayes, and B. C. Reed (2008), The Vegetation Drought Response Index (VegDRI): A new integrated approach for monitoring drought stress in vegetation, *GISci. Remote Sens.*, *45*, 16–46.

Daly, C., R. P. Neilson, and D. L. Phillips (1994), A statistical-topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteorol.*, *33*, 140–158.

Hayes, M., M. Svoboda, N. Wall, and M. Widhalm (2011), The Lincoln declaration on drought indices: Universal meteorological drought index recommended, *Bull. Am. Meteorol. Soc.*, *92*, 485–488.

Heim, R. R., Jr. (2002), A review of twentieth-century drought indices used in the United States, *Bull. Am. Meteorol. Soc.*, *83*, 1149–1165.

Houborg, R., M. Rodell, B. Li, R. Reichle, and B. Zaitchik (2012), Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations, *Water. Resour. Res.*, *48*, W07525, doi:10.1029/2011WR011291.

Ingber, L. (1989), Very fast simulated re-annealing, *Math. Comput. Model.*, *12*, 967–993.

Karl, T. R. (1986), The sensitivity of the Palmer Drought Severity Index and Palmer's Z-index to their calibration coefficients including potential evapotranspiration, *J. Climate Appl. Meteorol.*, *25*, 77–86.

Livneh, B., Y. Xia, K. E. Mitchell, M. B. Ek, and D. P. Lettenmaier (2010), Noah LSM snow model diagnostics and enhancements, *J. Hydrometeorol.*, *11*, 721–738.

McKee, T. B., N. J. Doesken, and J. Kleist (1993), The relationship of drought frequency and duration to time scales, in *Preprints, 8th Conference on Applied Climatology, 17–22 January*, edited by C. A. Anaheim, pp. 179–184, American Meteorological Society, Boston, MA.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, *21*, 1087–1092.

Mitchell, K. E., et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.*, *109*, D07S90, doi:10.1029/2003JD003823.

Mo, K. C. (2008), Model-based drought indices over the United States, *J. Hydrometeorol.*, *9*, 1212–1230.

Mo, K. C., and M. Chelliah (2006), The modified Palmer Drought Severity Index based on the NCEP North American Regional Reanalysis, *J. Appl. Meteorol. Climatol.*, *45*, 1362–1375.

Mo, K. C., L.-C. Chen, S. Shukla, T. J. Bohn, and D. P. Lettenmaier (2012), Uncertainties in North American Land Data Assimilation Systems over the Contiguous United States, *J. Hydrometeorol.*, *13*, 996–1009.

Nash, J. E., and J. V. Sutcliffe (1970), River flow for forecasting through conceptual models part 1—A discussion of principle, *J. Hydrol.*, *10*, 282–290.

Sen, M. K., and P. L. Stoffa (1996), Bayesian inference Gibbs' sampler and uncertainty estimation in geophysical inversion, *Geophys. Prospect.*, *44*, 313–350.

Sheffield, J., Y. Xia, L. Luo, E. F. Wood, M. Ek, and K. E. Mitchell (2012), Drought Monitoring with the North American Land Data Assimilation System (NLDAS): A framework for merging model and satellite data for improved drought monitoring, in *Remote Sensing of Drought: Innovative Monitoring Approaches*, edited by B. Wardlow, M. Anderson, and J. Verdin, p. 270, Taylor and Francis, London, U. K.

Shukla, S., and A. Wood (2008), Use of a standardized runoff index for characterizing hydrologic drought, *Geophys. Res. Lett.*, *35*, L02405, doi:10.1029/2007GL032487.

Svoboda, M., et al. (2002), The drought monitor, *Bull. Am. Meteorol. Soc.*, *83*, 1181–1190.

Troy, T. J., E. F. Wood, and J. Sheffield (2008), An efficient calibration method for continental-scale land surface modeling, *Water Resour. Res.*, *44*, W09411, doi:10.1029/2007WR006513.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno (2010), A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index, *J. Clim.*, *23*, 1696–1718.

Wang, A., T. J. Bohn, S. P. Mahanama, R. D. Koster, and D. P. Lettenmaier (2009), Multimodel ensemble reconstruction of drought over the continental United States, *J. Clim.*, *22*, 2694–2712.

Wei, H., Y. Xia, K. E. Mitchell, and M. B. Ek (2013), Improvement of the Noah land surface model for warm season processes: Evaluation of water and energy flux simulation, *Hydrol. Processes*, *27*, 297–303.

Xia, Y., M. K. Sen, C. S. Jackson, and P. L. Stoffa (2004), Multidataset study of optimal parameter and uncertainty estimation of a land surface model with Bayesian stochastic inversion and multicriteria method, *J. Appl. Meteorol.*, *43*, 1477–1497.

Xia, Y., et al. (2012a), Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *J. Geophys. Res.*, *117*, D03109, doi:10.1029/2011JD016048.

Xia, Y., et al. (2012b), Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow, *J. Geophys. Res.*, *117*, D03110, doi:10.1029/2011JD016051.

Xia, Y., M. B. Ek, D. Mocko, C. Peters-Lidard, J. Sheffield, J. Dong, and E. F. Wood (2013), Uncertainties, correlations, and optimal blends of drought indices from the NLDAS multiple land surface model ensemble, *J. Hydrometeorol.*, doi:10.1175/JHM-D-13-058.1.

Xia, Y., J. Sheffield, M. B. Ek, J. Dong, N. Chaney, H. Wei, J. Meng, and E. F. Wood (2014), Evaluation of multi-model simulated soil moisture in NLDAS-2, *J. Hydrol.*, *512*, 107–125, doi:10.1016/j.jhydrol.2014.02.027.